Fast online identification of power system dynamic behavior

Panagiotis N. Papadopoulos, Member, IEEE  
Jovica V. Milanović, Fellow, IEEE  
School of Electrical and Electronic Engineering  
The University of Manchester  
Manchester, UK  
panagiotis.papadopoulos@manchester.ac.uk,  
jovica.milanovic@manchester.ac.uk

Pratyasa Bhui, Student Member, IEEE  
Nilanjan Senroy, Member, IEEE  
Department of Electrical Engineering  
Indian Institute of Technology Delhi  
New Delhi, India  
pratyasansec@gmail.com,  
nsenroy@ee.iitd.ac.in

Abstract—This paper discusses the methodology for fast prediction of power system dynamic behavior. A combination of features that can be obtained from PMU data is proposed, that can improve the prediction time while keeping high accuracy of prediction. Several combinations of features including generator rotor angles, kinetic energy, acceleration and energy margin are used to train and test decision trees for the online identification of unstable generator groups. The predictor importance for trained decision trees is also calculated to highlight in more detail the effect of using different predictors.

Index Terms—decision trees, dynamic security assessment, online transient stability.

I. INTRODUCTION

In recent years, due to technical economic and environmental reasons, there is an ever increasing trend to utilize power system assets more efficiently. This means that power systems may be driven to be operated closer to stability limits. For this reason, the close to real time identification of power system dynamic behavior becomes important, since it can enable the application of corrective control actions. Impeding instabilities, and in the worst case scenario blackouts, could be avoided by using fast online methods for the identification of power system dynamic behavior.

For many corrective control actions, including generator tripping, load shedding and controlled islanding, it is essential to predict the dynamic behavior of the system fast, to both improve the efficiency of the applied corrective action and to increase the number of options for available corrective control actions.

Recently machine learning techniques have been applied to tackle the prediction of online identification of power system dynamic behavior. Decision Trees (DTs) [1]-[6], Ensemble Decision Trees (EDTs) [7], Support Vector Machine (SVM) [8], [9] and Artificial Neural Networks (ANNs) [10] are some of the most common methods used in online dynamic security assessment. In most of the cases, the prediction focuses on whether the system will remain stable or not (binary classification) [1]-[5], [7]-[9]. Additionally, in [6], [10] and [11] the resulting groups of unstable generators are also identified.

For machine learning methods, the required features are based on measured responses, usually obtained from Phasor Measurement Unit (PMU) data. Generator rotor angles are commonly used features since they can reveal important information for the dynamic behavior of generators. In [6], generator rotor angles are used as predictors by multiclass classifiers (DTs) for the online identification of unstable generator groups. Longer duration of rotor angle responses used by the machine learning methods can improve the accuracy of the prediction since more information are obtained for the system dynamic behavior as time progresses.

This paper focuses on improving the prediction time, by making it even faster, for online identification of power system dynamic behavior methods, while keeping high accuracy. Additional features such as energy margin, normalized kinetic energy of generators, total kinetic energy and generator acceleration are investigated. All chosen features can be calculated from PMU measurements and their calculation is very fast. The features are used as predictors to train DTs for the online identification of unstable generator groups. The features refer to both the entire system (e.g. energy margin, total kinetic energy) or to individual generators and it is therefore expected to help in both improving the performance related to distinguishing between stable and unstable cases as well as in identifying specific unstable generator groups.

II. METHODOLOGY

Dynamic time domain simulations are initially performed to generate the datasets used to train DTs following the approach described below. PMU measurement data are afterwards used online from the DTs to predict the dynamic behavior of the system, i.e. predict whether the system exhibits instability and identify the unstable generator groups.
Simulated data are used in this paper instead of PMU measurements to illustrate the methodology. In practical applications the rotor angles can be obtained from PMU measurements either following the electrical calculation method (as indicated in Annex F in IEEE Standard for Synchrophasor Measurements for Power Systems [12]) or even assuming that rotor angles are directly measured from generator rotors. Alternatively, if available, dynamic state estimators that utilize PMU measurement data can be used, as they could offer higher accuracy [13]. The effect of PMU measurement errors and possible signal loss on the performance of DTs for the online identification of power system dynamic behaviour have been investigated in [14] and [15]. Using ensemble algorithms, such as the C5.0 boosting algorithm used in this paper, reduces the possible impact of PMU measurement errors on the accuracy of the identification. Different features are tested to improve the performance of DTs, especially for very fast prediction time. In this paper, the terms predictors and features of the DTs (both denoted as $X_F$) are used interchangeably.

Transient stability of the system after a fault depends primarily on a) distribution of kinetic energy among generators after fault clearing, b) total kinetic energy after fault clearing, c) maximum potential energy absorbing ability of post-fault network. Generator rotor angles (as they evolve in time) and acceleration after the fault is cleared, can also provide valuable information considering the system dynamic behavior and especially considering individual generator dynamic behavior. Therefore, combinations of the above mentioned measures are used as predictors for DTs.

A. Decision Trees For Online Identification

Hierarchical clustering is initially applied on the generator rotor angle values to define unstable generator groups, as in [6]. The Euclidean distance is used to measure the similarity between clusters and the cutoff value to form the clusters is chosen as 360 degrees. This way the generators are categorized in groups where at least one generator of each group has 360 degrees difference from at least one generator of the other groups.

The identified unstable generator grouping patterns are afterwards used as the targets to train DTs as multiclass classifiers in a similar manner to [6]. The predictors related to generator rotor angles are measurement samples $k=1…n_{DR}$ of the rotor angles $\delta_k$ corresponding to the time duration from the instance the fault is cleared until $t_{DF}$ seconds for a number of $i=1…m$ generators in the system. There are $n_{DR}$ samples of $\delta_k$ for each one of the $m$ generators, so in total there are $n_{DF} \cdot m$ predictors related to generator rotor angles.

Using only generator rotor angles $\delta_k$ for duration $t_{DF}$, has been shown to provide good accuracy [6]. The accuracy of DTs increases as the number of samples from rotor angles increases, at the expense of having to wait longer (increasing $t_{DF}$) to predict the impending unstable generator groups. When corrective control actions are concerned, reducing the decision time $t_{DF}$ is important, since it will increase the effectiveness of the applied corrective actions. The value of $t_{DF}$ can be therefore defined as a trade-off between the accuracy of the prediction and the time delay in corrective control actions.

In this paper, the time $t_{DF}$ is reduced while keeping the accuracy high, by adding additional features that can be calculated very fast and that provide information for the whole system as well as for specific generators. The C5.0 boosting algorithm is used in this paper for training DTs, following the finding that it performs well for this class of problems as demonstrated in [6].

B. Feature Based on Transient Energy Function- Energy Margin

Energy function for power system transient stability studies, shown in (1), is well established [17] and used for contingency ranking and other transient stability studies. It describes the total system transient energy for the post disturbance system.

$$V(\omega_i, \theta_i) = V_{KE} + V_{PE} = 0.5 M_{eq} \omega_i^2 \sum_{i=1}^{m} P_i(\theta_i - \theta_i^f) + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} [C_{ij}(\cos \theta_i^f \cos \theta_j^f) + D_{ij}(\sin \theta_i^f - \sin \theta_j^f)]$$

(1)

where, $\omega_i$ is the speed of the $i^\text{th}$ generator, $M_{eq}$ is the moment of inertia of two machine equivalent, $\omega_i$ is the speed of two machine equivalent, $E$ is the internal emf, $P_i=P_{me} E_i^2 G_i$, $\theta_i$ is the post fault stable equilibrium point, $m$ is the number of generators, $C_{ij}=E_i E_j B_{ij}$ and $D_{ij}=E_i E_j G_{ij}$. The first term in (1) is the kinetic energy and depends on generator speeds only. Last three terms together form the potential energy of the system which depends only on generator angles.

If energy at fault clearing $V_{cl}$ is higher than the maximum potential energy that the post disturbance power system can absorb (this point is called the Controlling Unstable Equilibrium Point – CUEP, $\theta^*$), then the system will be unstable after fault clearing and vice versa. Therefore the difference of energy at fault clearing ($V_{cl}$) and the CUEP ($V_{cl}$) is used as an index to predict transient stability. If $\Delta V$, defined in (2), is negative at fault clearing ($\theta=\theta_{cl}$), the system will become unstable.

$$\Delta V(\omega_i, \theta_i) = \Delta V_{KE} + \Delta V_{PE} = -0.5 M_{eq} \omega_i^2 \sum_{i=1}^{m} P_i(\theta_i - \theta_i^f) - \sum_{i=1}^{m} \sum_{j=i+1}^{m} C_{ij}(\cos \theta_i^f \cos \theta_j^f) + D_{ij}(\sin \theta_i^f - \sin \theta_j^f)$$

(2)

This feature is expected to mainly improve the binary classification ability of DTs, i.e. distinguishing between stable and unstable cases and therefore reduce the errors related to either stable cases classified as unstable and vice versa. It is not expected to significantly help in identifying the specific unstable generator pattern.

C. Features Based on Normalized Kinetic Energy

Normalized Kinetic Energy (KE) is calculated as the kinetic energy gained by the generators divided by the total kinetic energy gained within 0.1s after the fault occurrence. For a fault at location $L$, the normalized kinetic energy is the
individual generator \( ke_{il} \) divided by total kinetic energy, as shown in (3), [18]. An instant of 0.1s is chosen for the above calculation as it is a typical fault clearing time. This feature has many advantages as it does not change for different fault types and only slightly varies with fault clearing time [18]. It is expected to mainly improve the multiclass classification performance of the DTs, since it provides information related to individual generators. The Total Kinetic Energy (TKE) gained by the generators during the fault can also be used as an additional feature and is expected to help in classification between stable and unstable cases.

\[
KE_{IL} = \left[ ke_{IL}, ke_{2L}, \ldots, ke_{mL} \right] = \sum_{i=1}^{m} ke_{il}
\]

Thus, \( \sum_{i=1}^{m} ke_{il} = 1 \)

\[D. \text{ Features based on generator acceleration}\]

Acceleration can be calculated from measured responses of generator rotor angles using (4), (5). The first instance of acceleration (namely \( A_i \), for \( i = 1 \ldots m \) generators) after the fault is cleared for each generator is used as a feature in this paper.

\[
\omega(t) = \frac{\delta(t) - \delta(t - \Delta t)}{\Delta t}
\]

\[
a(t) = \frac{\omega(t) - \omega(t - \Delta t)}{\Delta t}
\]

\[E. \text{ Predictor Importance - Sensitivity analysis}\]

The vector containing all the possible features for one observation is shown in (6) for \( m \) generators. The features included in (6) refer to five distinct categories, i.e. generator Rotor Angles (RA), generator Kinetic Energy (KE), Total Kinetic Energy (TKE), Energy Margin (EM) and Acceleration (A). Different combinations of the features shown in (6) are tested to identify the impact on the accuracy of DTs.

\[
X_f = [\delta_{11} \ldots \delta_{1m} \ldots \delta_{mDF} \ m \ ke_{1} \ ke_{2} \ldots \ ke_{m} \ \sum_{i=1}^{m} ke_{i} \ \Delta V A_{1} \ldots A_{m}]
\]

After a DT is trained, the importance of each predictor used can be calculated using a sensitivity measure \( S_i \), as defined in (7). The predictor importance \( VI_i \), can be then computed as the normalized sensitivity, as shown in (8) following a one-at-a-time approach [19], [20].

\[
S_j = \frac{V_j}{V(Y)}
\]

\[
VI_j = \frac{S_j}{\sum_{i=1}^{p} S_i}
\]

where \( V \) is the output variance considering all predictors, \( V_j \) is the variance without considering predictor \( X_j \) and \( p \) is the number of predictors used. As described above, for some of the features there are several individual predictors that correspond to each feature category, i.e. there are \( m_{DF} \) m predictors related to RA, \( m \) predictors related to KE, \( m \) predictors related to A, one predictor for EM and one for TKE. For each trained DT, the sensitivity measure for predictors related to the above measures can be added, in order to calculate the overall importance of each feature category, namely \( S_{RA}, S_{KE}, S_{A}, S_{EM} \) and \( S_{TKE} \). Since the provided features in (6) are not necessarily all used to generate the resulting splitting rules of the DT after the training is finished, calculating the predictor importance can highlight the impact of the predictors utilized in the decision process.

\[F. \text{ Accuracy of DTs}\]

To evaluate the accuracy of DTs in more detail, several error measures, given as percentages are used. Errors \( E_1 \), \( E_2 \) and \( E_3 \) are used as defined in (9)-(11). \( E_1 \) is related to the number of cases that are stable and misclassified as unstable, \( E_2 \) to the number of cases that are unstable and are misclassified as stable and \( E_3 \) to the number of cases that are classified as an incorrect unstable generator grouping pattern. The overall error can be calculated by adding \( E_1 + E_2 + E_3 \). The overall accuracy of the DTs presented in Section III is therefore related to the errors through 100-(\( E_1 + E_2 + E_3 \)).

\[
E_1 = \frac{\text{no of stable cases misclassified unstable}}{\text{no of all cases}} \times 100 \%
\]

\[
E_2 = \frac{\text{no of unstable cases misclassified stable}}{\text{no of all cases}} \times 100 \%
\]

\[
E_3 = \frac{\text{no of cases with misclassified pattern}}{\text{no of all cases}} \times 100 \%
\]

III. RESULTS

Five types of DTs are trained and tested using different feature combinations ranging from only using RA to using a combination of RA, EM, KE, TKE and A as explained in Section II.

\[A. \text{ IEEE 39 Bus Test Network}\]

The IEEE 39 Bus test network is used in this study implemented in PSS/E software [16]. For simplicity of presentation and to focus the discussion on the effects of different features used in the study (generator rotor angles, kinetic energy, acceleration and energy margin), only one, the base loading condition of the test system, is used in this paper. All generators are modeled using classical models. The training database for the DTs consists of 234 simulated responses for three-phase self-clearing faults at each bus of the test network. For each bus, the fault duration varies from 0.05s up to 0.3s with a step of 0.05s. Therefore, 6 faults are simulated for each one of the 39 buses of the test network. A similar testing database with 234 contingencies is generated as above, with fault durations varying from 0.07s up to 0.32s with a step of 0.05s.

\[B. \text{ Unstable generator groups}\]

From the application of hierarchical clustering, 56 different unstable grouping patterns are observed in the test network for both the training and testing dataset. 19 of those patterns are included only in the test dataset, which means that these 19 patterns have not been observed during the training of the DTs and are therefore guaranteed to be misclassified.

\[C. \text{ DT overall accuracy}\]

In Fig. 2 the performance of the DTs on the test dataset is presented for different prediction times, i.e. using observations
for certain time duration to calculate the features. The performance is evaluated by observing the accuracy of DTs (i.e. the % of cases that the correct unstable generator grouping pattern is identified) for different prediction times. Therefore, improvement in performance can be related to either faster prediction time keeping the same accuracy, or increasing the accuracy for longer prediction time. It should be mentioned that KE, TKE, EM and A are calculated during the fault or during the first instances after the fault is cleared (as described in Section II), and therefore remain constant irrespective of prediction times.

![Figure 1. IEEE 39 bus test network.](image)

In the case where only rotor angles are used, the accuracy is below 86% for prediction times up to 300 ms and increases above 90% for 400 ms prediction time. When adding EM, KE, TKE and A as features, there is an increase in the accuracy of DTs for prediction times less than 400 ms ranging from 2.5% to 4%, compared to the case when only rotor angles are used. For prediction times larger than 400 ms however, there is no significant increase in the accuracy. In the case when only EM, KE and TKE or EM, KE, TKE and A are used as features (rotor angles are not used), the accuracy is high even for very short prediction times in the order of 50 ms. Including rotor angles as features along with EM, KE, TKE and A improves the accuracy only for prediction times higher than 200ms. Comparing between the cases when no rotor angles are used as predictors, using only EM and KE exhibits low accuracy. Adding TKE improves by more than 4% the accuracy and adding A, a further 1.5%.

![Figure 2. Accuracy of DTs for different prediction time using different features.](image)

### D. Investigation of different types of errors

To highlight in more detail the impact of different features on the accuracy of DTs, the three errors defined in Section II F are presented in Fig. 3. In all cases, the error related to stable cases identified as unstable (E1) is low. The error related to unstable cases identified as stable (E2) is relatively higher than E1. Comparing between the different DTs, E2 is highest for the case when only EM, KE and TKE are used as predictors. Moreover, there is a slight increase in E2 when rotor angles with a longer duration are used (compared to shorter duration). This can be attributed to the fact that when a large duration of rotor angles is used, more weight is given by the DT in identifying between the different classes of instability as shown by the significant decrease of error E3 (when using long duration of rotor angles) and also from the sensitivity analysis presented in the next part of the paper. Considering the error of misclassifying the specific generator grouping pattern, it is the highest of all errors in general. Comparing between the different DTs, E3 is higher in the case of including only EM, KE and TKE and in the case when RA (50 ms), EM, KE, TKE and A are used as predictors. Including a short duration of rotor angles can shift the focus of the actual predictors used from other important features, such as the acceleration, as also shown in the sensitivity analysis in the following part. However, including a larger duration of rotor angles (600 ms) can reduce significantly error E3.

![Figure 3. Errors for DTs trained using different features.](image)

### E. Predictor Importance - Sensitivity analysis results

The predictor importance can provide information on the actual use of predictors to derive splitting rules, after the DT has been trained. A representative case for one of the trained DTs (RA-50ms+EM+KE+TKE+A) is shown in Table I where information about the actual predictors from which the DT splitting rules are derived, are provided. For three of the feature categories (i.e. RA, KE and A), the predictors related to specific generators can also be identified, since they are generator specific. For the rest of the feature categories (i.e. EM and TKE), there is always only one predictor, since they are related to the entire system.

The predictor importance related to different feature categories (i.e. rotor angles, energy margin, kinetic energy, total kinetic energy and acceleration) is shown in Fig. 4 for DTs trained with different features. In the case when RA (600 ms), EM, KE, TKE and A are used as features, the most important feature category is the rotor angles (15 predictors related to RA with a total of 16 predictors). The energy margin is also used but none of the other features are included in the decision rules of the obtained DT. When the same feature set as above but shorter time duration of rotor angles are used, all the feature categories are included in the decision rules, with rotor angles still having the largest importance, followed by
acceleration. This suggests that the use of a short duration of rotor angles is not sufficient and additional features are needed to improve the accuracy. In the case of EM, KE, TKE and A the acceleration becomes the most important feature category while in the case of EM, KE and TKE it shifts to kinetic energy. The ability of the last two mentioned DTs to perform well considering error $E_3$ presented above (slightly better than the case when short duration of rotor angles is also included as a feature) can be explained by this shift in the predictor importance. The final predictors that are actually used to derive the splitting rules of the DT can affect the obtained accuracy. Moreover, since both of the above mentioned DTs (EM+KE+TKE and EM+KE+TKE+A) have similar $E_3$ error, it can be concluded that both the generator acceleration and kinetic energy can be used to distinguish between specific unstable generator groups, with acceleration providing a slight additional advantage.

### Table 1. Predictor Importance Analysis for DT with Features RA(50ms), EM, KE, TKE and A

<table>
<thead>
<tr>
<th>Feature Category</th>
<th>Number of predictors</th>
<th>Related generator</th>
<th>Aggregated Predictor Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>5</td>
<td>4, 6, 7, 8, 9</td>
<td>0.373</td>
</tr>
<tr>
<td>EM</td>
<td>1</td>
<td>-</td>
<td>0.089</td>
</tr>
<tr>
<td>KE</td>
<td>2</td>
<td>5, 9</td>
<td>0.15</td>
</tr>
<tr>
<td>TKE</td>
<td>1</td>
<td>-</td>
<td>0.084</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>1, 6, 7, 9</td>
<td>0.304</td>
</tr>
</tbody>
</table>

Figure 4. Predictor importance for DTs trained using different features.

### IV. Conclusions

In this paper, the use of various features for training DTs to predict online unstable generator groups is investigated. Features based on both time domain simulations and direct methods are used that refer to the entire system behavior as well as to specific generators. The main aim is to reduce the decision time while still keeping a high prediction accuracy.

The generator kinetic energy with the energy margin and acceleration can improve the prediction accuracy for very fast prediction times (even as low as 50 ms) up to 4%. For higher prediction times (larger than 200 ms), including the rotor angles of generators further improves the accuracy. Therefore, the proposed added features can improve the performance in two ways: i) by increasing the accuracy for very fast prediction times (even when no rotor angles are used) and ii) by increasing the accuracy for slower prediction time in combination with rotor angles. A more detailed evaluation of the classification errors as well as the predictor importance is also carried out. The results highlight that the use of rotor angles for very short duration (50 ms) might impact negatively the accuracy of DTs since they might shift the focus of the DT to different features. When long duration of rotor angles is used (600 ms), the need of including additional features is not very significant. For fast prediction times with high accuracy a combination of energy margin, kinetic energy and generator acceleration is suggested to be used.

### References


